Distributed Resource Management in Multi-hop Cognitive Radio Networks for Delay Sensitive Transmission

Hsien-Po Shiang and Mihaela van der Schaar

Department of Electrical Engineering (EE), University of California Los Angeles (UCLA) Los Angeles, CA
{hpshiang, mihaela} @ee.ucla.edu

ABSTRACT

In this paper, we investigate the problem of multi-user resource management in multi-hop cognitive radio networks for delay-sensitive applications. Since the tolerable delay does not allow propagating global information back and forth throughout the multi-hop network to a centralized decision maker, the source nodes and relays need to adapt their actions (transmission frequency channel and route selections) in a distributed manner, based on *local* network information. We propose a distributed resource management algorithm that allows network nodes to exchange information and that explicitly considers the delays and cost of exchanging the network information over the multi-hop cognitive radio networks. The term "cognitive" refers in our paper to both the capability of the network nodes to achieve large spectral efficiencies by dynamically exploiting available frequency channels *as well as* their ability to learn the "environment" (the actions of interfering nodes) based on the designed information exchange. Note that the node competition is due to the mutual interference of neighboring nodes using the same frequency channel. Based on this, we adopt a multi-agent learning approach, adaptive fictitious play, which uses the available interference information. We also discuss the tradeoff between the cost of the required information exchange and the learning efficiency. The results show that our distributed resource management approach improves the PSNR of multiple video streams by more than 3dB as opposed to the state-of-the-art dynamic frequency channel/route selection approaches without learning capability, when the network resources are limited.

Index Terms: distributed resource management, cognitive radio networks, multi-hop wireless networks, multi-agent learning, delay sensitive applications.

I. INTRODUCTION

The demand for wireless spectrum has increased and will keep increasing rapidly in the foreseeable future with the introduction of multimedia applications such as YouTube, peer to peer multimedia networks, and distributed gaming. However, scanning through the radio spectrum reveals its inefficient occupancy in most

This work was supported by ONR.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **2008**	2. REPORT TYPE	3. DATES COVERED **00-00-2008 to 00-00-2008**

4. TITLE AND SUBTITLE **Distributed Resource Management in Multi-hop Cognitive Radio Networks for Delay Sensitive Transmission**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of California Los Angeles (UCLA),Department of Electrical Engineering (EE),Los Angeles,CA,90095**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
In this paper, we investigate the problem of multi-user resource management in multi-hop cognitive radio networks for delay-sensitive applications. Since the tolerable delay does not allow propagating global information back and forth throughout the multi-hop network to a centralized decision maker, the source nodes and relays need to adapt their actions (transmission frequency channel and route selections) in a distributed manner, based on local network information. We propose a distributed resource management algorithm that allows network nodes to exchange information and that explicitly considers the delays and cost of exchanging the network information over the multi-hop cognitive radio networks. The term "cognitive" refers in our paper to both the capability of the network nodes to achieve large spectral efficiencies by dynamically exploiting available frequency channels as well as their ability to learn the "environment" (the actions of interfering nodes) based on the designed information exchange. Note that the node competition is due to the mutual interference of neighboring nodes using the same frequency channel. Based on this, we adopt a multi-agent learning approach, adaptive fictitious play, which uses the available interference information. We also discuss the tradeoff between the cost of the required information exchange and the learning efficiency. The results show that our distributed resource management approach improves the PSNR of multiple video streams by more than 3dB as opposed to the state-of-the-art dynamic frequency channel/route selection approaches without learning capability, when the network resources are limited.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**35**	

frequency channels. Hence, the Federal Communications Commission (FCC) suggested in 2002 [1] improvements for spectrum usage, which enable more efficient allocations of frequency channels to license-exempt users without impacting the primary licensees. Based on this, cognitive radio networks [2][3] were proposed which enable wireless users to *sense* and *learn* the surrounding environment and correspondingly *adapt* their transmission strategies.

In such cognitive wireless environments, two main challenges arise. The first challenge is how to sense the spectrum and model the behavior of the primary licensees to identify available frequency channels (spectrum holes)[1]. The second challenge is how to manage the available spectrum resources among the license-exempt users to satisfy their QoS requirements while limiting the interference to the primary licensees. In this paper, we focus on the second problem, i.e. the resource management, and rely on the existing literature for the first challenge [4][5].

The majority of the resource management research in cognitive radio networks has focused on a single-hop wireless infrastructure [6]-[10]. In this paper, we focus on the resource management problem in the more general setting of multi-hop cognitive radio networks. A key advantage of such flexible multi-hop infrastructures is that the same infrastructure can be re-used and reconfigured to relay the content gathered by various transmitting users (e.g. sources nodes) to their receiving users (e.g. sinks nodes). These users may have different goals (application utilities etc.) and may be located at various locations. For the multi-hop infrastructure, there are three key differences as opposed to the single-hop case. First, the users have as available network resources not only the vacant frequency channels (spectrum holes or spectrum opportunities [2][6]) as in the single-hop case, but also the routes through the various network relays to the destination nodes. Second, the transmission strategies will need to be adapted not only at the source nodes, but also at the network relay nodes. In cognitive radio networks, network nodes are generally capable of sensing the spectrum and modeling the behavior of the primary users and thereby, identifying the available spectrum holes. In multi-hop cognitive radio networks, the network nodes will also need to model the behavior of the other neighbor nodes (i.e. other secondary users) in order to successfully optimize the routing decisions. In other words, network relays also require a learning capability in the multi-hop cognitive radio network. Third, to learn and efficiently adapt their decisions over

[1] In the wireless environment without primary licensees, such as the ISM band, there is no such problem. The main challenge is the resource management problem.

time, the wireless nodes need to possess accurate (timely) information about the channel conditions, interference patterns and other nodes transmission strategies. However, in a distributed setting such as a multi-hop cognitive radio network, the information is decentralized, and thus, there is a certain delay associated with gathering the necessary information from the various network nodes. Hence, an effective solution for multi-hop cognitive radio networks will need to tradeoff the “value” of having information about other nodes versus the transmission overheads associated with gathering this information in a timely fashion across different hops, in terms the utility impact.

In this paper, we aim at learning the behaviors of interacting cognitive radio nodes that use simple interference graph (similar to the spectrum holes used in [6][8]) to sequentially adjust and optimize their transmission strategies. We apply a multi-agent learning algorithm – the fictitious play (FP) [15] to model the behavior of neighbor nodes based on the information exchange among the network nodes. We focus on delay-sensitive applications such as real-time multimedia streaming, i.e. the receiving users need to get the transmitted information within a certain delay. Due to the informationally decentralized nature of the multi-hop wireless networks, a centralized resource management solution for these delay-constrained applications is not practical [14], since the tolerable delay does not allow propagating information back and forth throughout the network to a centralized decision maker. Moreover, the complexity and the information overhead of the centralized optimization grow exponentially with the size of the network. The problem is further complicated by the dynamic competition for wireless resources (spectrum) among the various wireless nodes (i.e. source nodes/relays). The centralized optimization will require a large amount of time to process and the collected information will no longer be accurate by the time transmission decisions need to be made. Hence, a distributed resource management solution, which explicitly considers the availability of information, the transmission overheads and incurred delays, as well as the value of this information in terms of the utility impact is necessary.

The paper is organized as follows. In Section II, we discuss the main challenges of the dynamic resource management in multi-hop cognitive radio networks and the related works. Section III provides the multi-hop cognitive radio network settings and strategies and Section IV gives problem formulation of the distributed resource management for delay sensitive transmission in such networks. In Section V, we determine how to quantify the rewards and costs associated with various information exchanges in the multi-hop cognitive radio

networks. In Section VI, we propose our distributed resource management algorithms with the information exchange and introduce the adopted multi-agent learning approach – adaptive fictitious play in the proposed algorithms. Simulation results are in Section VII. Finally, Section VIII concludes the paper.

II. Main Challenges and Related Works

A. Main challenges in multi-hop cognitive radio networks

To design such a distributed resource management in multi-hop cognitive radio networks, several main challenges need to be addressed:

- **Dynamic adaptation to a time-varying network environment**

Multi-hop cognitive radio networks are generally experiencing the following dynamics: 1) the primary users directly affect the spectrum opportunities available for the secondary users, 2) the mobility of the network relays that affects the network topology, 3) the traffic load variation due to multiple applications simultaneously sharing the same network infrastructure, and 4) the time-varying wireless channel conditions. Given the dynamic nature of the cognitive radio networks, wireless nodes need to learn, dynamically self-organize and strategically adapt their transmission strategies to the available resources without interfering the primary licensees. Due to these time-varying dynamics, the outcomes of these interactions do not need to converge to an equilibrium, i.e., disequilibrium and perpetual adaptation of strategies may persist, as long as the performance of the delay sensitive application is maximized [15]. Hence, repeated information exchange among network nodes is required for nodes to efficiently learn and keep adapting to the changing network dynamics.

- **Information availability in multi-hop infrastructures**

Due to the informationally-decentralized nature of the multi-hop infrastructure, the exchanged network information is only useful when it can be conveyed in time. The timeliness constraint of the information exchange depends on the delay deadline of the applications, the information overhead, and the condition of the network links, etc. Hence, the value of information in terms of its impact on the users' utilities will need to be quantified for the different settings of the multi-hop cognitive radio network. This information will impact the accuracy with which the wireless nodes can model the behavior of other nodes (including the primary users) and hence, the efficiency with which they can respond to this environment by adequately optimizing their transmission strategies.

B. Related works

Distributed dynamic spectrum allocation is an important issue in cognitive radio networks. Various approaches have been proposed in recent years. In [8], a decentralized cognitive MAC protocols are proposed based on the theory of Partially Observable Markov Decision Process (POMDP), where a secondary user is able to model the primary users through Markovian state transition probabilities. In [9], the authors investigated a game-theoretic spectrum sharing approach, where the primary users are willing to share spectrum and provide a determined pricing function to the secondary users. In [10], a no-regret learning approach is proposed for dynamic spectrum access in cognitive radio networks. However, these studies focus on dynamic spectrum management for the single-hop network case.

Exploiting frequency diversity in wireless multi-hop networks has attracted enormous interests in recent years. In [11], the authors propose a distributed allocation scheme of sub-carriers and power levels in an orthogonal frequency-division multiple-access-based (OFDMA) wireless mesh networks. They proposed a fair scheduling scheme that hierarchically decouples the sub-carrier and power allocation problem based on the limited local information that is available at each node. In [12], the authors focus on the distributed channel and routing assignment in heterogeneous multi-radio, multi-channel, multi-hop wireless networks. The proposed protocol coordinates the channel and route selection at each node, based on the information exchanged among two-hop neighbor nodes. However, these studies are not suitable for cognitive radio networks, since they ignore the dynamic nature of spectrum opportunities and users (network nodes) need to estimate the behavior of the primary users for coexistence. To the best of our knowledge, the dynamic resource management problem in multi-hop cognitive radio networks has not been addressed in literature.

In summary, the paper makes the following contributions.

a) We propose a dynamic resource management scheme in multi-hop cognitive radio network settings based on periodic information exchange among network nodes. Our approach allows each network nodes (secondary users and relays) to exchange their spectrum opportunity information and select the optimal channel and next relay to transmit delay sensitive packets.

b) We investigate the impact of the information exchange collected from various hops on the performance of the distributed resource management scheme. We introduce the notion of an "information cell" to explicitly identify

the network nodes that can convey timely information. Importantly, we investigate the case that the information cell does not cover all the interfering neighbor nodes in the interference graph.

c) The proposed dynamic resource management algorithm applies FP [15], which allows various nodes to learn their spectrum opportunity from the information exchange and adapt their transmission strategies autonomously, in a distributed manner. Moreover, we discuss the tradeoffs between the cost of the required information exchange and the learning efficiency of the multi-agent learning approach in terms of the utility impact.

Next, we present our network settings of the multi-hop cognitive radio networks.

III. Multi-hop Cognitive Radio Networks - Settings and Strategies

A. *Network entities*

In this paper, we assume that a multi-hop cognitive radio network involves the following network entities and their interactions:

- **Primary Users (PUs)** are the incumbent devices that possess transmission licenses for specific frequency bands (channels). Without loss of generality, we assume that there are M frequency channels in the considered cognitive radio network. We also assume that the maximum number of primary users that can be present in the network equals M. Note that these primary users can only occupy their assigned (licensed) frequency channels and not other primary users' channels. Since the primary users are licensed users, they will be guaranteed an interference-free environment [2][4]. When a primary user is not transmitting data using its assigned frequency channel, a spectrum hole is formed at the corresponding frequency channel.
- **Secondary Users (SUs)** are the autonomous wireless stations that perform channel sensing and access the existing spectrum holes in order to transmit their data. The secondary users can occupy the spectrum holes available in the various frequency channels. In this paper, the secondary users are deploying delay sensitive applications. Specifically, we assume that there are V delay sensitive applications simultaneously sharing the cognitive radio network infrastructure, having unique source and destination nodes. These secondary users are able to deploy their applications across various frequency channels and routes.
- **Network Relays (NRs)** are autonomous wireless nodes that perform channel sensing and access the existing spectrum holes in order to relay the received data to one of its neighboring nodes or SUs. Hence, unlike in the SUs case, there is no source or destination present at the NRs. Note that multiple applications can use

the same NR using different frequency channels.

B. Source traffic characteristics

Let V_i denote the delay sensitive application of the i-th SU. Assume that the application V_i consists of packets in K_i priority classes. The total number of applications is V. We assume that there are a total of $K = \sum_{i=1}^{V} K_i + 1$ priority classes (i.e., $\mathbf{C} = \{C_1,\ldots,C_K\}$). The reason for adding an additional priority class is because the highest priority class C_1 is reserved for the traffic of the primary users. The rest of the classes $C_k, k > 1$ can be characterized by:

- λ_k, the impact factor of a class C_k. For example, this factor can be obtained based on the money paid by a user (different service levels can be assigned for different SUs by the cognitive radio network), based on the distortion impact experienced by the application of each SU or based on the tolerated delay assigned by the applications. The classes of the delay sensitive applications are then prioritized based on this impact factor, such that $\lambda_k \geq \lambda_{k'}$ if $k < k', k = 2,\ldots,K$. The impact factor is encapsulated in the header (e.g. RTP header) of each packet.
- D_k, the delay deadline of the packets in a class C_k. In this paper, a packet is regarded useful for the delay sensitive applications only when it is received before its delay deadline.
- L_k, the average packet length in the class C_k.

A variety of delay sensitive applications can use the cognitive radio set-up discussed in this paper. Multimedia transmission such as video streaming or video conferencing can be examples of such applications [14]. We assume in this paper that an application layer scheduler is implemented at each network node to send the most important packet first based on the impact factor encapsulated in the packet header.

C. Multi-hop cognitive radio network specification

We consider a multi-hop cognitive radio network, which is characterized by a general topology graph $\mathcal{G}(\mathbf{M},\mathbf{N},\mathbf{E})$ that has a set of primary users $\mathbf{M} = \{m_1,\ldots,m_M\}$, a set of network nodes $\mathbf{N} = \{n_1,\ldots,n_N\}$ (include SUs and NRs) and a set of network edges (links) $\mathbf{E} = \{e_1,\ldots,e_L\}$ (connecting the SUs and NRs). There are a total of N nodes and L links in this network. Each of these N network nodes is either a secondary user (as a source or a destination node) or a network relay.

We assume that $\mathbf{F} = \{f_1, ..., f_M\}$ is the set of frequency channels in the network, where M is the total number of the frequency channels. To avoid interference to the primary users, the network nodes can only use spectrum holes for transmission. Hence, to establish a link with its neighbor nodes, each network node $n \in \mathbf{N}$ can only use the available frequency channels in a set $\mathbf{F}_n \subseteq \mathbf{F}$. Note that these wireless nodes in a cognitive radio network will continuously sense the environment and exchange information and hence, $\mathbf{F}_n$ may change over time depending on whether the primary users are transmitting in their assigned frequency channels.

The network resource for a network node $n \in \mathbf{N}$ of the multi-hop cognitive radio network includes the routes composed by the various links and frequency channels. We define the resource matrix $\mathbf{R}_n = [R_{ij}] \in \{0,1\}^{L \times M}$ for the network node n as follows:

$$R_{ij} = \begin{cases} 1, & \text{if link } e_i \text{ is connected to the node } n \\ & \text{and the frequency channel } f_j \text{ is available.} \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Whether or not the resource R_{ij} is available to node $n \in \mathbf{N}$ depends not only on the topology connectivity, but also on the interference from other traffic using the same frequency channel. Next, we discuss the interference from other users (including the primary users).

D. *Interference characterization*

Recall that the highest priority class C_1 is always reserved in each frequency channel for the traffic of the primary users. The traffic of the SUs can be categorized into $K-1$ priority classes ($C_2, ..., C_K$) for accessing frequency channels. The traffic priority determines its ability of accessing the frequency channel. Primary users in the highest priority class C_1 can always access their corresponding channels at any time. The traffic of the SUs can only access the spectrum holes for transmission. Hence, we define two types of interference to the secondary users in the considered multi-hop cognitive radio network:

1) Interference from primary users.

In practical cognitive networks, even though primary users have the highest priority, secondary users will cause some level of interference to the primary users due to their imperfect awareness (sensing) of the primary users. The primary users' interference depends on the location of the M primary users. We rely on methods such as in [5] that consider the power and location of the secondary users to ensure that the

secondary users do not exceed some critical interference level to the primary users. We also assume that the spectrum opportunity map is available to the secondary users as in [6][10]. Since the primary users will block all the neighbor links using its frequency channel, a network node n will sense the channel and obtain the Spectrum Opportunity Matrix (SOM) of the primary users:

$$\mathbf{Z}_n = [Z_{ij}] \in \{0,1\}^{L\times M},\ \text{with}\ Z_{ij} = \begin{cases} 1, & \text{if the primary user is occupying frequency channel } f_j \\ & \text{and the link } e_i \text{ can interfere with the primary user.} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

A simple example is illustrated in Figure 1, which indicates the SOM of the primary users and the resource matrix of each network node in the multi-hop cognitive radio network.

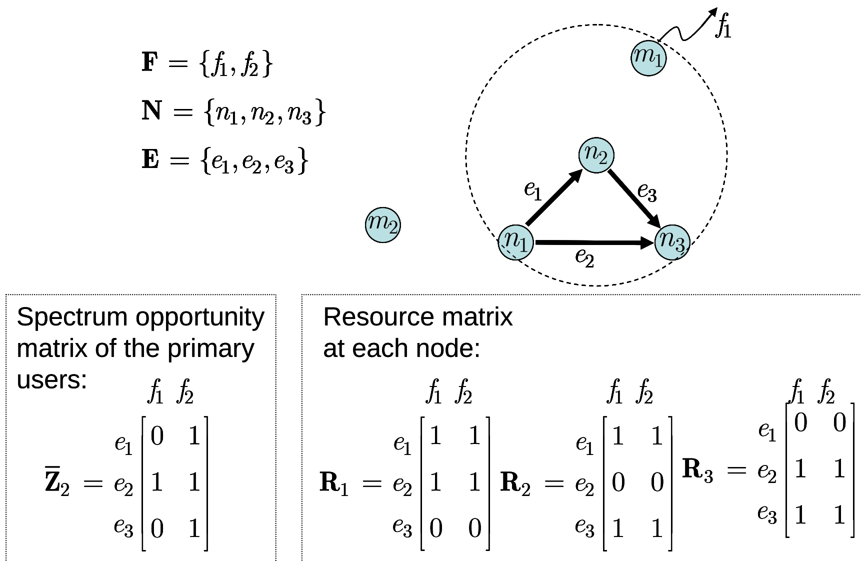

Fig. 1. A simple multi-hop cognitive radio network with three nodes and two frequency channels.

2) Interference from competing secondary users.

We define $\mathbf{I}_k = [I_{ij}] \in \{0,1\}^{L\times M}$ as the Interference Matrix (IM) for the traffic in priority class $C_k, k \geq 2$.

$$I_{ij} = \begin{cases} 1, & \text{if link } e_i \text{ using frequency channel } f_j \text{ can be} \\ & \text{interfered by the traffic of priority class } C_k. \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The interference caused by the traffic in priority class C_k can be determined based on the interference graph of the nodes that transmit the traffic (as in [10]). The interference graph is defined as the corresponding links that are interfered by the transmission of the class C_k traffic[2]. The IM can be computed by the information

---

[2] In a wireless environment, the transmission of *neighbor links* can interfere with each other and significantly impact their effective transmission time. Hence, the action of a node can impact and be impact by the action of the other relay nodes. In order to coordinate these neighboring nodes, we construct the interference matrix with binary “1” and “0”.

exchange among the neighbor nodes.

The available resource matrix can be masked out by the SOM and IM of the higher priority classes, i.e. $\mathbf{R}_{nk}^{(I)} = \mathbf{R}_n \otimes \overline{\mathbf{I}}_{k-1} \otimes ... \otimes \overline{\mathbf{Z}}_n$, where the notation $\otimes$ represents element-wise multiplication of the matrixes and $\overline{\mathbf{I}}$ denotes the inverse operation, which turns 1 into 0 and 0 into 1. The resulting resource matrix $\mathbf{R}_{nk}^{(I)}$ represents the *available resource* around the network node n for the class C_k traffic under the interference of other higher priority traffic (classes). Next, we define the actions available to the network nodes in a multi-hop cognitive radio network.

E. Nodes' actions

We define the action of the network node n in order to relay the delay sensitive application V_i as $A_n = (e \in \mathbf{E}_n, f \in \mathbf{F}_n)$. We assume that a network relay n can select a set of links to its neighbor nodes (links connected to node n) $\mathbf{E}_n \subseteq \mathbf{E}$. Corresponding to the actions, we define the transmission strategy vector of the network node n as $\mathrm{s}_n = [s_A \mid A = (e \in \mathbf{E}_n, f \in \mathbf{F}_n)]$, where s_A represent the probability that the network node n will choose an action A. We refer to an action at a node n as a "*feasible action*" for transmitting a class C_k traffic, if $A = (e,f)$ is an "*available resource*" in $\mathbf{R}_{nk}^{(I)}$ (i.e. element $R_{ef} = 1$ in $\mathbf{R}_{nk}^{(I)}$), since in this case the selected link and frequency channel do not interfere with the traffic in the higher priority classes. That is,

$$\hat{\mathbf{A}}_n(k) = \{A = (e,f) \mid \mathbf{R}_{nk}^{(I)} = [R_{ef}]^{L\times M}, R_{ef} = 1\}\,. \tag{4}$$

We denote the set of all the feasible actions for node n as $\hat{\mathbf{A}}_n(k)$ for class C_k traffic. We next determine the corresponding delay based on different actions, which considers the deployed cross-layer transmission strategies in order to compute the Effective Transmission Time (ETT) [19] over the transmission links.

Each network node n computes the ETT $ETT_{nk}(e,f),\ \text{with}\ e \in \mathbf{E}_n, f \in \mathbf{F}_n$ for transmitting delay sensitive applications in priority class C_k:

$$ETT_{nk}(e,f) = \frac{L_k}{T_n(e,f) \times (1 - p_n(e,f))}\,. \tag{5}$$

$T_n(e,f)$ and $p_n(e,f)$ represent the transmission rate and the packet error rate of the network node n using the frequency channel f over the link e. $T_n(e,f)$ and $p_n(e,f)$ can be estimated by the MAC/PHY layer link

adaptation [20]. Specifically, we assume that the channel condition of each link-frequency channel pair can be modeled using a continuous-time Markov chain [17] with a finite number of states $\mathbf{S}_{(e,f)}^n$. The time a channel condition spends in state $i \in \mathbf{S}_{(e,f)}^n$ is exponentially distributed with parameter ν_i (rate of transition at state i in transitions/sec). We assume that the maximum transition rate[3] of the network is ν and the variation of the channel conditions in a time interval $\tau \le 1/\nu$ is regarded negligible.

Define the action vector $\mathbf{A}_i = [A_n \mid n \in \boldsymbol{\sigma}_i]$ as the vector of the actions of all the network relay nodes for transmitting V_i. Assume that the ith delay sensitive application V_i are transmitted from the source node $n_i^s \in \mathbf{N}$ to the destination node $n_i^d \in \mathbf{N}$ with a total of q_i packets. The routes of V_i are denoted as $\boldsymbol{\sigma}_i = \{\sigma_{ij} \mid j = 1,...,q_i\}$, where σ_{ij} is the route of the jth packet in V_i. A route σ_{ij} is a set of link-frequency pairs that the packets flow through, i.e.

$$\sigma_{ij} = \{(e,f) \mid \text{ the } j\text{th packet of } V_i \text{ flows through link } e \text{ using frequency channel } f\}. \tag{6}$$

Note that if the action of a certain relay node changes, the corresponding route $\sigma_{ij}(\mathbf{A}_i)$ of relaying V_i also changes. We denote the end-to-end delay of the packets transmitted using the route $\sigma_{ij}(\mathbf{A}_i)$ as $d_{ij}(\sigma_{ij}(\mathbf{A}_i))$. Based on the topology, each network relay node receiving a packet can decide where to relay the packet to and using which frequency channel, in order to minimize its end-to-end delay $d_{ij}(\sigma_{ij}(\mathbf{A}_i))$. Finally, to calculate $d_{ij}(\sigma_{ij}(\mathbf{A}_i))$, the source node need to obtain the delay information from other nodes according to the actions taken by the relay nodes, i.e.

$$d_{ij}(\sigma_{ij}(\mathbf{A}_i)) = \sum_{n \in \sigma_{ij}} ETT_{nk}(\mathbf{A}_i), \text{ for } V_i \in C_k \,. \tag{7}$$

IV. Resource Management Problem Formulation over Multi-hop Cognitive Radio Networks

By examining the cumulated ETT values, the objective of a delay sensitive application is to minimize its own end-to-end packet delay. The centralized and proposed distributed problem formulations are subsequently

[3] In case that some of the channel conditions change severely in the network, a threshold ν_{th} can be set by protocols to avoid these fast-changing nodes and the ν is hence selected as the maximum transition rate below this threshold value.

provided.

- **Centralized problem formulation with global information available at the sources**

If we assume that the global information[4] $\mathcal{G}_i$ is available to the source node n_i^s for the delay sensitive application V_i, the route $\sigma_{ij}(\mathbf{A}_i, \mathcal{G}_i)$ can be determined for each packet j of V_i. The centralized optimization can be performed at every source node in order to maximize the utility u_i. Hence, for application V_i we have:

$$\begin{aligned} \mathbf{A}_i^{opt} &= \arg\max u_i(\mathbf{A}_i, \mathcal{G}_i) \\ &\text{subject to } A \in \hat{\mathbf{A}}_n \text{ for all } A \in \mathbf{A}_i \end{aligned}, \tag{8}$$

$$\text{where } u_i(\mathbf{A}_i, \mathcal{G}_i) = \sum_{j=1}^{q_i} \lambda_{ij} \cdot \text{Prob}\{d_{ij}(\sigma_{ij}(\mathbf{A}_i, \mathcal{G}_i)) \le D_{ij}\}, \quad D_{ij} = D_k \text{ and } \lambda_{ij} = \lambda_k \text{ if } j \in C_k. \tag{9}$$

However, due to the limited wireless network resource, the end-to-end delay constraint $d_{ij}(\sigma_{ij}(\mathbf{A}_i, \mathcal{G}_i)) \le D_k$ can make the optimization solution infeasible. Hence, a sub-optimal greedy algorithms that perform optimizations sequentially from the highest priority class to the lowest priority class are commonly adopted [25][14]. Specifically, for class C_k, the following optimization is considered:

$$\begin{aligned} \mathbf{A}_{ik}^{opt} &= \arg\min \sum_{j \in C_k} d_{ij}(\sigma_{ij}(\mathbf{A}_{ik}, \mathcal{G}_i)) \\ &\text{subject to } d_{ij}(\sigma_{ij}(\mathbf{A}_{ik}, \mathcal{G}_i)) \le D_k, \quad , \\ &\qquad A \in \hat{\mathbf{A}}_n \text{ for all } A \in \mathbf{A}_{ik}. \end{aligned} \tag{10}$$

where $\mathbf{A}_{ik} = [A_n \mid n \in \boldsymbol{\sigma}_{ij}, j \in C_k]$.

Due to the informationally decentralized nature of the multi-hop wireless networks, the centralized solution is not practical for the multi-user delay sensitive applications, as the tolerable delay does not allow propagating the global information $\mathcal{G}_i$ back and forth throughout the network to a centralized decision maker. For instance, the optimal solution depends on the delay d_{ij} incurred by the various packets across the hops, which cannot be timely relayed to a source node. For instance, when the network environment is time-varying, the gathered global information $\mathcal{G}_i$ can be inaccurate due to the propagation delay for this information. Moreover, the complexity of the centralized optimization grows exponentially with the number of classes and nodes in the network. The problem is further complicated by the dynamic adaptation of the transmission strategies deployed

---

[4] The word "global information" means the information gathered from every node throughout the network. We discuss the required information in Section V.

by the wireless nodes, which impacts their spectrum access and hence, implicitly, the performance of their neighbor nodes. The optimization will require a large amount of time to process and the collected information might no longer be accurate by the time transmission decisions need to be made.

In summary, in the studied dynamic cognitive radio network, the decisions on how to adapt the aforementioned actions at sources and relays need to be performed in a distributed manner due to these informational constraints. Hence, a "decomposition" of the optimization problem into distributed strategic adaptation based on the available local information is necessary.

- **Proposed distributed problem formulation with local information at each node:**

Instead of gathering the entire global information $\mathcal{G}_i$ at each source, we propose a distributed suboptimal solution that collects the local information $\mathcal{L}_n$ at node n to minimize the expected delay of the various applications sharing the same multi-hop wireless infrastructure. Note that at each node n, the end-to-end delay for sending a packet $j \in C_k$ in equation (10) can be decomposed as:

$$d_{ij}(\sigma_{ij}) = d_n^P(\sigma_{ij}) + E[\tilde{d}_n(k,\sigma_{ij})], \tag{11}$$

where $d_n^P(\sigma_{ij})$ represents the past delay that packet j has experienced before it arrives at node n and $E[\tilde{d}_n(k,\sigma_{ij})]$ represents the expected delay from the node n to the destination of the packet $j \in C_k$. The sending packet $j \in C_k$ is determined by the application layer scheduler according to the impact factor λ_k. The information about λ_k can be encapsulated in the packet header and $d_n^P(\sigma_{ij})$ can be calculated based on the timestamp available in the packet header. The priority scheduler at each node ensures that the higher priority classes are not influenced by the lower priority classes (see equation (10)). Since at the node n the value of $d_n^P(\sigma_{ij})$ is fixed, the optimization problem at the node n becomes:

$$\begin{aligned} A_n^{opt} &= \arg\min E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))] \\ &\text{subject to } E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))] \le D_k - d_n^P(\sigma_{ij}) - \rho,\ j \in C_k\ , \\ &\qquad A_n \in \hat{\mathbf{A}}_n \end{aligned} \tag{12}$$

where $E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))]$ represents the expected delay from the relay node n to the destination of the packets in class C_k . ρ represents a guard interval such that the probability $\text{Prob}\{E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))] + d_n^P(\sigma_{ij}) > D_k\}$ is small (as in [22]). To estimate the expected delay

$E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))]$ in equation (12), each network node n maintains an estimated transmission delay $E[\tilde{d}_n(k)]$ from itself to the destination for each class of traffic using the Bellman-Ford shortest-delay routing algorithm [17]. We assume that each node n maintains and updates a delay vector $\mathbf{d}_n = [E[\tilde{d}_n(2)],...,E[\tilde{d}_n(K)]]$ (note that the first priority class is reserved for the primary users) with elements for each priority class. Each network node exchanges such information to its neighbor nodes and selects the best action A_n^{opt} for the highest priority packet in the buffer of the network node n. We will discuss the minimum-delay routing/channel selecting algorithm in Section VI. Note that a group of packets in the buffer of a node n can take the action A_n, since the action is determined based on local information $\mathcal{L}_n$. Since in the cognitive radio networks, the available channel is time-variant, the information needs to be timely conveyed to the network node for the distributed optimization. Compared to the centralized approach in equation (8), the distributed resource management in equation (12) can adapt better to the dynamic wireless environment by periodically gathering local information. Next, we discuss the distributed resource management with information constraints in more detail.

V. DISTRIBUTED RESOURCE MANAGEMENT WITH INFORMATION CONSTRAINTS

A. *Considered medium access control*

In this paper, we assume that the required local information $\mathcal{L}_n$ is exchanged using a designated coordination control channel similar to [13]. Such a coordination channel can be selected from the existing ISM bands, since there is no primary licensee in these bands to interfere with. The transmission is time slotted and the time slot structure of a node is provided in Figure 2. We denote the time slot duration as t_I. The action A_n are selected at each node, during each time slot, after the coordination interval (that includes the channel sensing for SOM and the information exchange for IM). We denote the coordination interval at the network node n as $d_I(\mathcal{L}_n)$. The goal of the coordination interval at each time slot is to provide the feasible action set $\hat{\mathbf{A}}_n$ for the channel access and the relay selection of the packet transmission. We will discuss how to obtain $\hat{\mathbf{A}}_n$ based on the SOM and the IM among the neighboring nodes when we introduce the proposed algorithm, in Section VI.

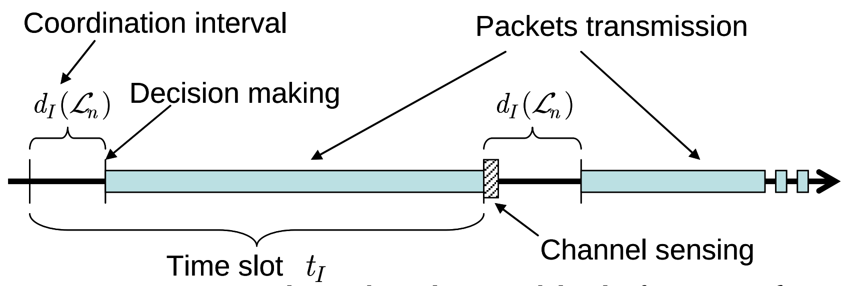

Fig. 2. Transmission time line at the node n with local information $\mathcal{L}_n$.

Besides the SOM and IM, the information required in the coordination interval should also include the delay vectors $\mathbf{d}_n$ and the control messages for RTS/CTS coordination [8][12]. Note that the local information $\mathcal{L}_n$ does not need to include all these information in each time slot (except the control messages). For example, the SOM and IM can be collected in a different period, depending on the sensing and information exchange mechanism. Hence, the coordination duration $d_I(\mathcal{L}_n)$ will vary for different time slots, which will be discussed in more detail in Section V.*C*. Next, we investigate the benefit of acquiring information from different h-hop neighbor nodes, which also affects the duration of the coordination interval $d_I(\mathcal{L}_n)$.

B. Benefit of acquiring information and information constraints

For the network node n, the local information $\mathcal{L}_n$ gathered from different network nodes has different impact on decreasing the objective function $E[\tilde{d}_n(k,\sigma_{ij}(A_n,\mathcal{L}_n))]$ in equation (12). Let $\mathcal{I}_n(x) = \{\mathbf{I}_k(n_x, A_{n_x}), A_{n_x}, \mathbf{d}_{n_x} \mid n_x \in \mathbf{N}_x^n\}$ denote the set of local information gathered from the neighbor nodes, which is x hops away from node n, where $\mathbf{N}_x^n$ represents a set of nodes that is x hops away from node n. We define $\mathcal{L}_n(x) = \{\mathcal{I}_n(l) \mid l = 1,...,x\}$ as the local information gathered from all of these neighbor nodes. Given the local information $\mathcal{L}_n(x)$, we define the optimal expected delay as $K_n(k,x) = E[\tilde{d}_n(k,\sigma_{ij}(A_n^{opt},\mathcal{L}_n(x)))]$. The larger x will has a smaller expected delay $K_n(k,x)$. The benefit (reward) of the information $\mathcal{I}_n(x)$ for the class C_k traffic is denoted as $J_n(k,\mathcal{I}_n(x))$. In a static network case, $J_n(k,\mathcal{I}_n(x))$ is defined as:

$$J_n(k,\mathcal{I}_n(x)) \triangleq K_n(k,x-1) - K_n(k,x), \text{ if } x > 1. \tag{13}$$

We define $J_n(k,\mathcal{I}_n(1)) = K_n(k,1)$ since $\mathcal{L}_n(1) = \mathcal{I}_n(1)$. The reward of information $J_n(k,\mathcal{I}_n(x))$ can be regarded as the benefit (decrease of the expected delay) in terms of the expected delay $E[\tilde{d}_n(k,\sigma_{ij})]$ if the information $\mathcal{I}_n(x)$ is received by node n. Note that the optimal expected delay $K_n(k,x)$, given the

information $\mathcal{L}_n(x)$:

$$K_n(k,x) = K_n(k,1) - \sum_{l=2}^{x} J_n(k,\mathcal{I}_n(l)). \tag{14}$$

Equation (14) states that the optimal expected delay is a decreasing function of x, meaning that smaller expected delays can be achieved as more information is gathered. The improvement is quantified by the reward of the information $J_n(k,\mathcal{I}_n(l))$. Here, we ignore the cost of exchanging such information, which will be defined in the next subsection. Figure 3 shows a simple illustrative example of reward of information at node n, which is five hops away from the destination node of class C_k traffic. The more information $\mathcal{I}_n(x)$ available from nodes that is x hops away, the smaller optimal expected delay $K_n(k,x)$ can be obtained.

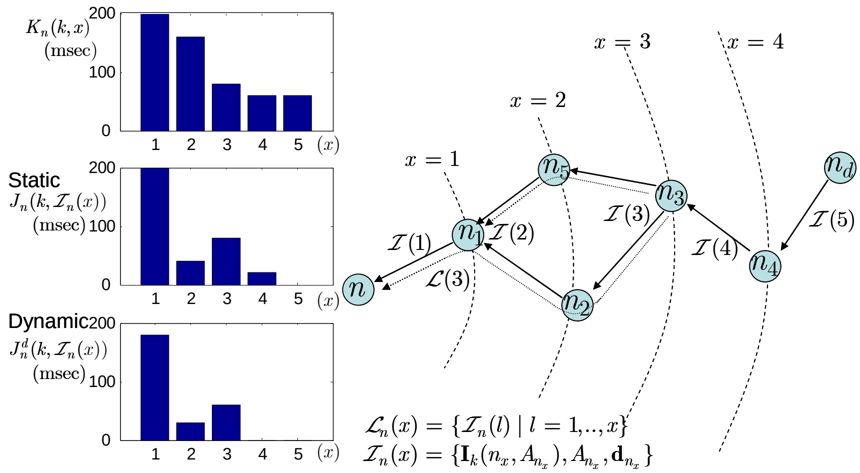

Fig. 3. Example of the static reward of information $J_n(k,\mathcal{I}_n(x))$, dynamic reward of information $J_n^d(k,\mathcal{I}_n(x))$ and optimal expected delay $K_n(k,x)$ (where the information horizon $h_n(k,\nu) = 3$, average packet length L_k =1000 bytes, and average transmission rate T = 6Mbps over the multi-hop network).

Let $\mathbf{J}_n(k) = [J_n(k,\mathcal{I}_n(x)),\ \text{for } 1 \le x \le H_n]$ denote the reward vector from 1-hop information to H_n-hop information, where $H_n = \max\{H_n^I, H_n^d\}$. H_n^d represents the shortest hop counts from the node n to the destination node of the class C_k traffic and H_n^I represents the interference range in terms of hop counts for node n. We also need to consider the hop count H_n^I in case that the destination node is close to the node n within the interference range. We assume that the reward vector $\mathbf{J}_n(k)$ is obtained when the network is first deployed and only updated infrequently, when SUs join or leave the network. Note that all the elements in $\mathbf{J}_n(k)$ are nonnegative, i.e. $J_n(k,\mathcal{I}_n(x)) \ge 0,\ \text{for } 1 \le x \le H_n$, due to the fact that knowing additional information cannot increase the expected delay $E[\tilde{d}_n(k,\sigma_{ij})]$ in a static network. However, if we consider the

propagation delay of such information exchange across the network in the dynamic network, the dynamic reward of information $J_n^d(k, \mathcal{I}_n(x))$ decreases as the hop count x increases. When the information of the further nodes reaches the decision node n, the information is more likely to be out-of-date (i.e. the information cannot reflect the exact network situation in a dynamic setting, since the network conditions and traffic characteristics are time-varying). Once the information is out-of-date, $J_n^d(k, \mathcal{I}_n(x)) = 0$, i.e. there is no benefit from gathering information that is out-of-date. Note that in a dynamic network, once $J_n^d(k, \mathcal{I}_n(x)) = 0$, $J_n^d(k, \mathcal{I}_n(x')) = 0$ for $x \le x' \le H_n$.

Therefore, in the dynamic network, we define the information horizon $h(k,\nu)$ such that

$$h_n(k,\nu) \triangleq \arg\max x \quad \text{subject to } J_n^d(k, \mathcal{I}_n(x)) > \phi(k,\nu), 1 \le x \le H_n . \tag{15}$$

where $\phi(k,\nu) \ge 0$ represents a minimum delay variation specified by the application which determines the minimum benefit of receiving local information for class C_k traffic. In fact, $h_n(k,\nu)$ depends on the variation speed ν of the wireless network condition (i.e. the transition rate of the Markovian channel condition model, see Section III.*E*). In a dynamic network with higher variation speeds ν (e.g. with high mobility), a higher threshold $\phi(k,\nu)$ is needed to guarantee that the information $\mathcal{I}_n(x)$ is still valuable and it should be exchanged. This results in a smaller information horizon $h_n(k,\nu)$. We illustrate this mobility issue in Section VII. Note that the information horizon $h_n(k,\nu)$ varies for different classes of traffic at different locations in the network. Since higher priority class traffic has more network resources than the lower priority class (i.e. they are scheduled first for optimization in equation (12)), the threshold value $\phi(k,\nu) \le \phi(k',\nu)$, if $k < k'$ and thereby, $h_n(k,\nu) \ge h_n(k',\nu)$, if $k < k'$. In other words, the information horizon $h_n(k,\nu)$ of a higher priority class C_k is larger than the information horizon $h_n(k',\nu)$ of a lower priority class $C_{k'}$.

Although the information horizon $h_n(k,\nu)$ can vary at different locations for different priority classes depending on the applications, the complexity of such implementation is high and the adaptation of the information horizon itself can be an interesting topic. Hence, we will leave the information horizon adaptation problem to our future research. For simplicity, we assume in this paper that the information horizon is only a function of the network variation speed ν, i.e. $h_n(k,\nu) = h(\nu)$. The information horizon $h(\nu)$ is determined

for the most important class among the SUs in the network. This definition of the information horizon $h(\nu)$ is aligned with [14], in which $h(\nu)$ is defined as the maximum number of hops that the information can be conveyed in τ, such that the network is considered unchanged (recall that any network changes within the interval $\tau(\nu) \leq 1/\nu$ can be regarded negligible).

Based on this information horizon $h(\nu)$, we assume that the network nodes within the $h(\nu)$ hops form an *information cell*. Only the local information $\mathcal{L}_n(h)$ *within* the information cell is useful to the node n, since the reward of information is zero, i.e. $J_n(k, \mathcal{I}_n(x)) = 0$ for $\forall x > h(\nu)$. In the dynamic network, network node n determines its action at time slot t based on the acquired information at the previous time slot $t-1$. The optimization problem in equation (12) can be written as:

$$\begin{aligned} A_n^{opt}(t) = \arg\min E[\tilde{d}_n(k, \sigma_{ij}(A_n, \mathcal{L}_n(h, t-1)))] \\ \text{subject to } E[\tilde{d}_n(k, \sigma_{ij}(A_n, \mathcal{L}_n(h, t-1)))] \leq D_k - d_n^P(\sigma_{ij}) - \rho,\ j \in C_k\ . \\ A_n \in \hat{\mathbf{A}}_n(t-1) \end{aligned} \tag{16}$$

Recall that the neighbor nodes of the node n are defined as the nodes that can interfere or can be interfered by the node n (within H_n^I hops), which may not align with the range of the information cell (within $h(\nu)$ hops). If all neighbor nodes are within the h-hop information cell, all necessary information are timely conveyed to the node n. Otherwise, the neighbor nodes that are too far away cannot convey the interference information to the node n in time. Since the required information cannot be acquired in time, the solution in equation (16) becomes suboptimal. We refer to this problem as “information exchange mismatch” problem.

Figure 4 illustrates two simple network examples with and without the mismatch problem. Note that in Figure 4(b), since the information cell does not cover all the interfering neighbor nodes, the center node n_2 will still be interfered by other secondary users. In fact, due to the nature of the multi-hop wireless environment, the network nodes that are far away from the node n have limited interference impact on node n_2. Hence, even though the information horizon h does not match the interference range, the performance degradation of the optimization problem in equation (16) using the local information $\mathcal{L}_n(h)$ is limited.

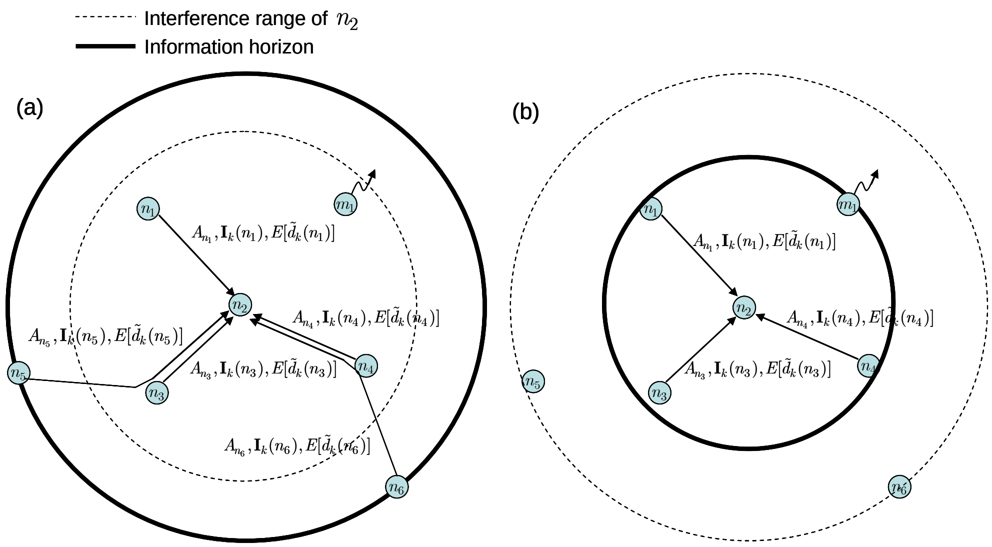

Fig 4. (a) 2-hop information cell network without information exchange mismatch problem.
(b) 1-hop information cell network with information exchange mismatch problem.

C. Cost of information exchange

In the previous subsection, we discuss the reward of information in an h-hop information cell while ignoring the negative impact of the information exchange. In this section, we discuss the cost (increase of the expected delay) due to this information exchange. Recall that the duration of the time slot is $t_I(\nu)$, which is also the interval between the repeated information exchanges in the network. We define there are c time slots in τ seconds, i.e.

$$t_I(\nu) = \frac{\tau(\nu)}{c}. \tag{17}$$

c defines the frequency of the decision making as well as the learning process, which will be discussed in detail in Section VI. Note that decisions can be made every t_I and this time slot duration is short enough compared to τ. Hence, the network changes in t_I is also negligible.

Recall that the coordination duration in a time slot for the network node n is $d_I(\mathcal{L}_n(h))$. Assume the information unit for the required information is $U^{(I)}$, $U^{(A)}$, and $U^{(d)}$ per class, respectively. Assume the average number of nodes in an h-hop information cell is $\bar{N}(h)$. The information time overhead of $\mathcal{L}_n(h)$ is on average $d_I(\mathcal{L}_n(h)) = \bar{N}(h)[(K-1)(U^{(d)} + U^{(I)}) + U^{(A)}]$.

Note that even though the information exchange is implemented in a designated coordination channel [13], a network node with a single antenna cannot transmit both the data and the control signals at the same time. This

information exchange time overhead decreases the effective transmission rate at node n using the line e and frequency channel f:

$$T_n'(e,f) = \frac{t_I(\nu) - d_I(\mathcal{L}_n(h))}{t_I(\nu)} \times T_n(e,f)\,. \tag{18}$$

Hence, the effective transmission time at a node n using the link e and frequency channel f to transmit a packet in class C_k becomes:

$$ETT_{nk}'(e,f) = \frac{t_I(\nu)}{t_I(\nu) - d_I(\mathcal{L}_n(h))} \times ETT_{nk}(e,f)\,. \tag{19}$$

In conclusion, the increase of the effective transmission time degrades the performance of the delay sensitive applications. The degradation depends on the content of the local information exchange $\mathcal{L}_n(h)$, and the network variation speed ν. Hence, the benefit $J_n^d(k, \mathcal{I}_n(x))$ in equation (15) will decrease due to this cost of the information. Hence, we denote the value of information with this cost consideration as $J_n^c(k, \mathcal{I}_n(x))$:

$$\begin{aligned} J_n^c(k, \mathcal{I}_n(x)) &= K_n'(k, x-1) - K_n'(k,x) \\ &= K_n(k, x-1) \times \frac{t_I(\nu)}{t_I(\nu) - d_I(\mathcal{L}_n(x-1))} - K_n(k,x) \times \frac{t_I(\nu)}{t_I(\nu) - d_I(\mathcal{L}_n(x))}\,. \end{aligned} \tag{20}$$

And the optimal information horizon $h_n(k,\nu)$ in equation (15) also decreases due to the cost. Next, we discuss the proposed distributed resource management algorithm based on the information exchanges and learning capabilities to tackle the optimization problem in equation (16).

VI. Distributed Resource Management Algorithms

Figure 5 provides a system diagram of the proposed distributed resource management. First, a packet $j \in C_k$ is selected from the application scheduler at the node n based on the impact factor λ_k of the packet and an action A_n is taken for that packet. The application layer information including C_k, L_k, D_k is conveyed to the network layer for this action decision. Network conditions $T_n(e,f), p_n(e,f)$ are then conveyed from the MAC/PHY layer for computing the ETT values using equation (5).

In addition to the $T_n(e,f), p_n(e,f)$, the action selection is impacted by the interference induced from the action of these neighbor nodes and hence, the information received from the neighbor nodes in the information cell. Recall that $\mathcal{L}_n(h) = \{\mathcal{I}_n(l) \mid l = 1,\ldots,h\}$. We use the notation $-n(h)$ to represent the set of the neighbor nodes

of the network node n in the h-hop information cell. Hence, the local information exchanged $\mathcal{L}_n(h) = \{\mathbf{I}_k(-n(h), A_{-n(h)}), A_{-n(h)}, \mathbf{d}_{-n(h)}\}$ across the network nodes is required. Hence, the node n knows the estimated delay $\mathbf{d}_{-n(h)}$ from its neighbor nodes to the destinations, so as the actions $A_{-n(h)}$ of its neighbor nodes and their IM $\mathbf{I}_k(-n(h), A_{-n(h)})$. Based on the delay information from the neighbor nodes $\mathbf{d}_{-n(h)}$, a network node can update its own estimated delay to the various destinations and determine the minimum-delay action based on Bellman-Ford algorithm [17].

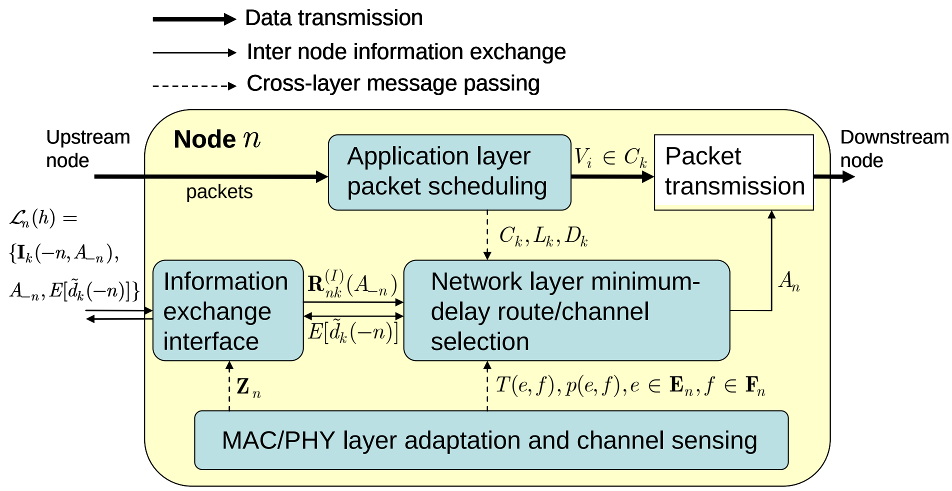

Fig. 5. System diagram of the proposed distributed resource management.

We separate the distributed resource management into two blocks at the node n as in Figure 5 – the information exchange interface block that regularly collects required local information and the route/channel selection block to determine the optimal action. We now discuss the role of the exchanged information and the two algorithms implemented in these blocks, respectively.

A. *Distributed resource management algorithms*

The next algorithm is performed at network node n at the information exchange interface in Figure 5.

Algorithm 1. Periodic information exchange algorithm:

Step 1. Collect the required information – the node n first collects the required information the SOM $\mathbf{Z}$ from channel sensing and $\mathcal{L}_n(h) = \{\mathbf{I}_k(-n(h), A_{-n(h)}), A_{-n(h)}, \mathbf{d}_{-n(h)}\}$ from the neighbor nodes in the information cell.

Step 2. Learn the behavior of the neighbor nodes – by continuously monitoring the actions of the neighbor nodes, node n can model the behavior of the neighbor nodes or learn a better transmission strategy using strategy vectors $\mathbf{s}(n') = [s_A(n') \mid A = (e \in \mathbf{E}_{n'}, f \in \mathbf{F}_{n'})]$, $n' \in -n(h)$, where $s_A(n')$ represents the probability

(strategy) of selecting an action A by the node n', which will be discussed in the next subsection.

Step 3. Estimate the resource matrix – from the SOM and the IM $\mathbf{I}_k(n', A_{n'})$ gathered from the neighbor node n', the resource matrix can be obtained for each class of traffic by $\mathbf{R}_{nk}^{(I)} = \mathbf{R}_n \otimes \overline{\mathbf{I}}_{k-1} \otimes ... \otimes \overline{\mathbf{Z}}_n$, which will be explained in Section VI.*A* in more details. Then the available resource $\mathbf{R}_{nk}^{(I)}(A_{-n})$ are provided to the network layer route/channel selection block stated in the Algorithm 2.

Step 4. Update information $\{\mathbf{I}_k(n, A_n), A_n, \mathbf{d}_n\}$ – based on the recently selected action A_n, the latest delay vector $\mathbf{d}_n$, and the IM $\mathbf{I}_k(n, A_n)$. Two types of interference model are considered in this paper when constructing the IM $\mathbf{I}_k(n, A_n)$ from equation (3):

1) A network node can transmit and receive packets at the same time – Note that a node cannot reuse a frequency channel $f \in \mathbf{F}_n$ used by its neighbor nodes. If a frequency channel is used by its neighbor nodes, all the elements in the column of the interference $\mathbf{I}_k(n, A_n)$ that is associated with the frequency channel are set to 1. Then the IM is exchanged to the nodes within the pre-determined information horizon h.
2) A network node cannot transmit and receive packets at the same time – In this case, if the frequency channel $f \in \mathbf{F}_n$ is used, all the elements in the column of the IM $\mathbf{I}_k(n, A_n)$ associated with the frequency channel are set to 1. In addition, if a network link $e \in \mathbf{E}_n$ is used by its neighbor nodes, all the elements of the IM $\mathbf{I}_k(n, A_n)$ that is associated with the node n are also set to 1, no matter what frequency channel it uses. Then the IM is exchanged to the nodes within the pre-determined information horizon h.

Step 5. Broadcast the information $\{\mathbf{I}_k(n, A_n), A_n, \mathbf{d}_n\}$ **and repeat the algorithm periodically in every** $t_I(\nu)$ **seconds.**

The next algorithm is performed at the network node n at the network layer minimum-delay route/channel selection block in Figure 5.

<u>*Algorithm 2. Minimum-delay route/channel selection algorithm:*</u>

Step 1. Determine the packet to transmit – based on the impact factor, one packet j in the buffer at the node n is scheduled to be transmitted. Assume the packet $j \in C_k$, and the information of C_k, L_k, $D_k - d_n^P$ are extracted or computed from the application layer.

Step 2. Construct the feasible action set – construct the feasible action set $\hat{\mathbf{A}}_n(k)$ from the resource matrix

$\mathbf{R}_{nk}^{(I)}$ given from the information exchange interface for the priority class C_k at the node n (see equation (4)).

Step 3. Estimate the channel condition – the transmission rate $T_n(e,f)$ and packet error rate $p_n(e,f)$ for each link-frequency channel pair $(e \in \mathbf{E}_n, f \in \mathbf{F}_n)$ are provided from the PHY/MAC layer through link adaptation [20].

Step 4. Calculate the expected delay toward the destination – for each action $A_n \in \hat{\mathbf{A}}_n(k)$ of the traffic class C_k:

$$E[\tilde{d}_n(k, A_n)] = ETT_{nk}(A_n) + E[\tilde{d}_{n'(A_n)}(k)], \text{ for } \forall A_n \in \hat{\mathbf{A}}_n(k), \tag{21}$$

where $E[\tilde{d}_{n'(A_n)}(k)]$ represents the corresponding element for the class C_k in the delay vector $\mathbf{d}_{-n}$ from the neighbor node $n'(A_n)$. $ETT_{nk}(A_n)$ can be calculated based on L_k, $T_n(e,f)$, and $p_n(e,f)$ using equation (5).

Step 5. Check the delay deadline – if $E[\tilde{d}_n(k)] \geq D_k - d_n^P - \rho$, drop the packet.

Step 6. Select the minimum delay action – if $E[\tilde{d}_n(k)] < D_k - d_n^P - \rho$, find the minimum-delay route and frequency channel selection, i.e. determine the optimal action A_n^{opt} from the feasible action set $\hat{\mathbf{A}}_n(k)$. In other words, the goal here is to solve equation (16) at node n:

$$A_n^{opt} = \arg \min_{A_n \in \hat{\mathbf{A}}_n(k)} E[\tilde{d}_n(k, A_n)]. \tag{22}$$

Note that the feasible action set $\hat{\mathbf{A}}_n(k)$ in equation (22) depends on the actions of other neighbor nodes A_{-n}. It is important for the network nodes to adopt learning approaches for modeling the behaviors of these network nodes to decrease the complexity of the dynamic adaptation. This will be discussed in the next subsection.

Step 7. Send RTS request – after determining the next relay and frequency channel, send RTS request indicating the determined action information A_n^{opt} to the next relay.

Step 8. Wait for CTS response and transmit the packets.

Step 9. Update the delay and the current action information – after selecting the optimal action, update the estimated delay $E[\tilde{d}_n(k)]$ using exponential moving average with a smoothing factor α:

$$E[\tilde{d}_n(k)] = \alpha \times E[\tilde{d}_n(k)]^{old} + (1-\alpha) \times E[\tilde{d}_n(k, A_n^{opt})], \tag{23}$$

and provide the updated delay vector $\mathbf{d}_n = [E[\tilde{d}_n(2)], ..., E[\tilde{d}_n(K)]]$ to Algorithm 1 at the information exchange interface. In Figure 6, we provide a block diagram of the proposed distributed resource management. For the

blocks that beyond the scope of this paper, we refer to [4][5] for channel sensing, [8][12] for RTS/CTS coordination, and [17] for the delay vectors.

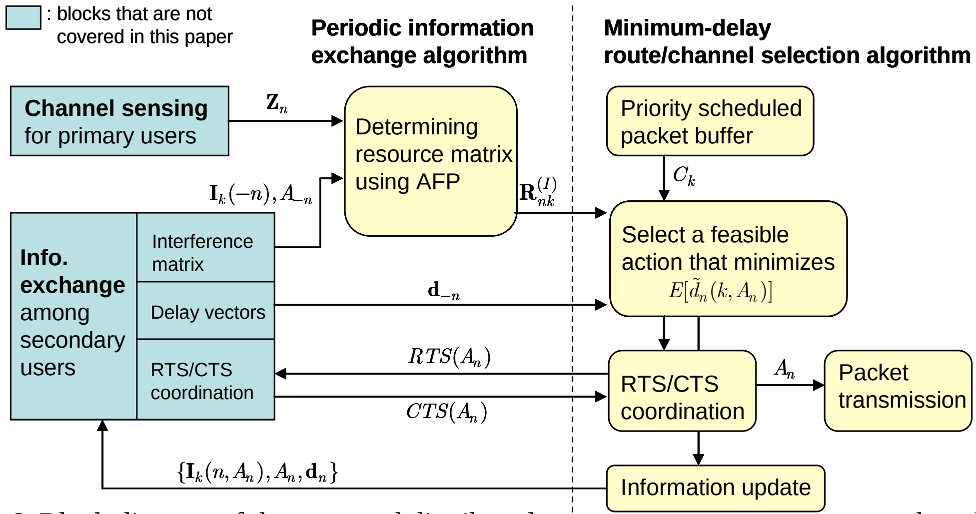

Fig. 6. Block diagram of the proposed distributed resource management at network node n .

B. *Adaptive fictitious play (AFP)*

We now provide a learning approach for the SUs to learn the feasible action set $\hat{\mathbf{A}}_n(k)$ in equation (22) for our distributed resource management algorithms. Specifically, based on the information exchange $\mathcal{L}_n(h)$, the behaviors of the neighbor nodes in the information cell can be learned (Step 2 of Algorithm 1) and based on the behaviors, the feasible action set $\hat{\mathbf{A}}_n(k)$ is determined. This motivates us to apply a well-known learning approach – fictitious play [15], applied when the SUs are willing[5] to reveal their current action information and thereby, they are able to model the behaviors (strategies) of other SUs (a model-based learning [18]). However, due to the information constraint discussed in the previous section, only the information from the neighbor nodes in the information cell is useful. Hence, we adapt the fictitious play learning approach to our considered network setting. Figure 7(a) provides a block diagram of the proposed distributed resource management algorithm using the adaptive fictitious play.

Note that only part of the SUs can be modeled via the learning approach depending on the information horizon. Specifically, a node n maintains a strategy vector over time $\mathbf{s}(n',t) = [s_A(n',t) \mid A = (e \in \mathbf{E}_{n'}, f \in \mathbf{F}_{n'})]$ for each of its neighbor nodes $n' \in -n(h)$ in the information cell.

---

[5] If the action information is not provided by the other secondary users, a node can learn its own strategy from its action payoffs – the estimated delay $E[\tilde{d}_n(k)]$. The learning approach refers to the reinforcement learning (a model-free learning [18]) or a payoff-based learning).

$s_A(n',t)$ represents the frequency selection strategy of the node n' making action A at time t, which is obtained using:

$$s_A(n',t) = \frac{r_A(n',t)}{\sum\limits_{A \in (\mathbf{E}_{n'}, \mathbf{F}_{n'})} r_A(n',t)}, \tag{24}$$

where $r_A(n',t)$ is the *propensity* [16] of node n' for taking action A at time t, which can be computed by:

$$r_A(n',t) = \alpha \times r_A(n',t-1) + I(A_{n'}(t) = A), \tag{25}$$

where $\alpha < 1$ is a discount factor quantifying the importance of the history value. $I(A_{n'}(t) = A)$ represents an indicator function such that,

$$I(A_{n'}(t) = A) = \begin{cases} 1, \text{ if the action of the node } n' \text{ at time } t \text{ is } A \\ 0, \text{ otherwise} \end{cases}. \tag{26}$$

Figure 7(b) shows how the network variation speed ν affects the size of the information cell and ultimately, the video performance. We will consider the mobility of the network relays to show this network variation impact in the next section.

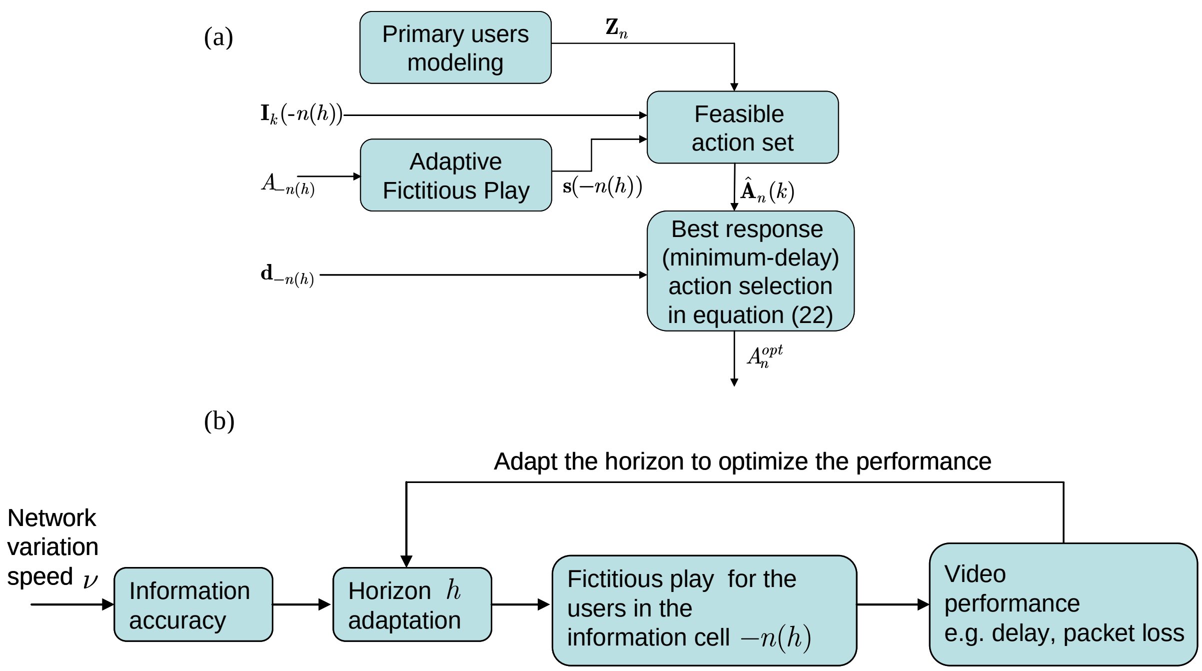

Fig. 7(a). Block diagram of the proposed distributed resource management algorithm using the AFP.
7(b). Impact of the network variation on the FP and the video performance.

As stated in Section III.*E*, $s_A(n',t)$ represent the probability that the network node n' will choose an action A. Hence, the probability $s_A(n',t)$ for modeling the node n' making an action A at time t will increase

with the actual times that the action A is selected. Based on the strategy $s_A(n',t)$, the adaptive fictitious play provides the estimated IM $\mathbf{I}_k$, and then the feasible action set $\hat{\mathbf{A}}_n(k)$ can be computed.

From the gathered IM $\mathbf{I}_k(n',A_{n'})$ from the neighbor node $n' \in -n(h)$, the node n can compute the expected IM from

$$\mathbf{I}_k^e = [I_{ij}^e] = \sum_{n' \in -n(k)} \mathbf{I}_k(n') = \sum_{n' \in -n(k)} \sum_A s_A(n') \mathbf{I}_k(n',A)\,. \tag{27}$$

Then, the node n can estimate the IM $\mathbf{I}_k$ for the traffic in class C_k:

$$\mathbf{I}_k = [I_{ij} \mid I_{ij} = \begin{cases} 1, \text{ if } I_{ij}^e \geq \mu \\ 0, \text{ if } I_{ij}^e < \mu \end{cases}], \tag{28}$$

where μ represents a threshold value that determines whether or not a link-frequency-channel pair (e,f) is considered to be occupied. Feasible action set $\hat{\mathbf{A}}_n(k)$ can hence be learned based on the resource matrix $\mathbf{R}_{nk}^{(I)} = \mathbf{R}_n \otimes \overline{\mathbf{I}}_{k-1} \otimes ... \otimes \overline{\mathbf{Z}}_n$ using equation (4). By learning the feasible action set $\hat{\mathbf{A}}_n(k)$, the best response actions are computed using equation (22).

C. Information exchange overhead reduction

The fictitious play suffers from a large information overhead, since it requires all the local information $\mathcal{L}_n(h) = \{\mathbf{I}_k(-n(h), A_{-n(h)}), A_{-n(h)}, \mathbf{d}_{-n(h)}\}$ in the h–hop information cell. From the cost of information exchange in equation (20), we know that the overhead can increase the expected delay, especially when the network changes slowly (i.e. with a large information cell). Hence, the overhead reduction is required to mitigate the performance degradation.

(1) Reducing the information horizon.

Recall that the information overhead of $\mathcal{L}_n(h)$ is $\bar{N}(h)[(K-1)(U^{(d)} + U^{(I)}) + U^{(A)}]$ in average ($\bar{N}(h)$ is the average number of nodes in an h-hop information cell). With an information horizon $h' < h$, the overhead becomes $\bar{N}(h')[(K-1)(U^{(d)} + U^{(I)}) + U^{(A)}]$, where $\bar{N}(h') < \bar{N}(h)$. Note that it is not always beneficial to decrease the overhead by reducing the information horizon. There exists a trade-off as discussed in Section V. The reward of information $J_n^d(k, \mathcal{I}_n(x))$, $x < h$ in equation (15) provides a metric to select the most valuable information from the nodes within the information cell.

(2) Reducing the number of classes.

From equation (12), we know that the higher priority classes will not be influenced by the lower priority classes. Hence, the information overhead can be reduced by ignoring the information exchange of the lower priority classes. The overhead becomes $\bar{N}(h)[(k'-1)(U^{(d)}+U^{(I)})+U^{(A)}]$, $k'<K$.

(3) Reducing the frequency of learning.

Although we divide c time slots in τ seconds, a network node n does not have to learn in all these c time slots. In other words, the periodic learning process of the node n does not have to be aligned with the information exchange (decision making). In order to avoid simultaneous learning among network neighbors in a distributed manner, at each time slot, the network node n updates the strategy vector $s_A(n',t)$ with probability $\varepsilon_n = b_n / c$ ($b_n \le c$), and keeps the same strategy vector with probability $1-\varepsilon_n$. In other words, the network node n chooses b_n time slots out of c time slots in τ seconds to model the behavior of other neighbor nodes. Note that the parameter b_n characterize the speed of learning at different network node n. The larger b_n gives the network node n faster learning capability. The information overhead of $\mathcal{L}_n(h)$ becomes $b_n / c \times \bar{N}(h)[(K-1)(U^{(d)}+U^{(I)})+U^{(A)}]$.

VII. Simulation Results

We simulate two video streaming applications that are transmitting videos V_1 "Coastguard" and V_2 "Mobile" (16 frames per GOP, frame rate of 30Hz, CIF format) over the same multi-hop cognitive radio network. Each video sequence is divided into four priority classes ($K_i = 4, K = 9$) with average packet length L_k = 1000 bytes and delay deadline D_k = 500 millisecond. Although the first priority class C_1 is reserved for the primary users, let us first consider the case when there are no primary users, i.e. only the SUs and NRs are transmitting. We assume that there are two frequency channels (M =2). The wireless network topology is shown in Figure 8 in a 100x100 meters region with N = 15 nodes and L = 22 links similar to the network settings in [21]. A link is established as long as the channel condition (described in the paper by the link SINR) is acceptable within the transmission distance (approximately 36 meters). Note that this transmission distance is not aligned with the interference range H_n^I. Neighbor nodes that are beyond the transmission distance can still interfere with each other.

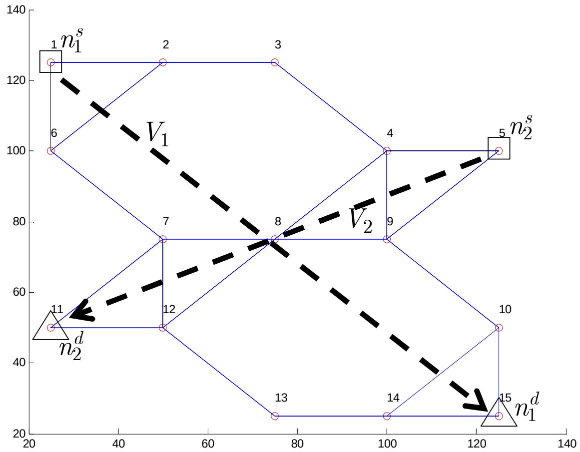

Fig. 8. Wireless network settings for the simulation of two video streams.

A. *The reward and cost of the information exchange*

First, we simulate the impact of the information including the reward J_n^d (see equation (13)) and cost J_n^c (see equation (20)) from the expected delay $E[\tilde{d}_n]$ using the adaptive fictitious play in Section VII with different information horizons. Figure 9 shows the resulting reward and cost of information at different locations for streaming video V_1 (at node n = 1, 7, and 13 on one of the routes of video V_1). The results show that a 1-hop information cell is enough when the interference range is 40 meters, since only the nodes that are 1 hop away can interfere with each other. If the interference range is 80 meters, the information exchange mismatch problem (see Section V) occurs and the appropriate information horizon for information exchange is then increased to 2.

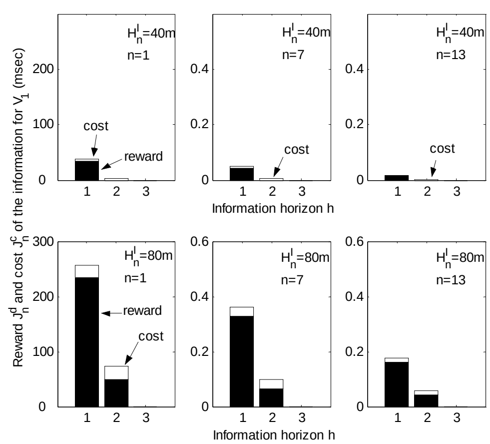

Fig.9. Reward J_n^d and cost J_n^c of different information horizon at different node for video V_1.

B. *Application layer performance with different information horizons and interference ranges*

We next compare the proposed dynamic resource management algorithm using adaptive fictitious play (AFP)

with two other resource management methods – AODV [23] with load balancing over the two available frequency channels (AODV/LB) and the Dynamic Least Interference Channel Selection [24] (DCS) extended to a network setting. Table I and II show the results of the Y-PSNR of the two video sequences using different approaches. The results show that the proposed algorithm using learning from the nodes within the information cell outperforms the alternative approaches. Especially, when the interference range is large (H_n^I = 80 meters), the proposed AFP approach significantly improves the video quality (X represents PSNR below 26 dB, which is unacceptable for a viewer).

For delay sensitive applications, we measure the packet loss rate (i.e. the probability that the end-to-end delay exceeds the delay deadline) for different approaches in Figure 10(a). The results of both applications are shown. The AODV represents the on-demand routing solution with only 1 frequency channel. The AODV/LB approach randomly distributes packets over the two available frequency channels. The DCS approach with cognitive ability selects a better frequency channel based on the link measurements and hence, improves the performance opposed to the AODV/LB. The AFP further improves the performance of both applications by learning the behaviors of the neighbor nodes. Interestingly, the benefit brought by the learning capability decreases as the network bandwidth increases. In other words, it is not worthy to be too intelligent in an environment with plenty of resource. Moreover, as shown in Figure 10(b), the improvement of 2-hop information cell is limited when the interference range is 40 meters. This is because the nodes that are two hops away have no impact on the current node and their information is not valuable (i.e. it does not impact the utility).

TABLE I.
THE CHARACTERISTIC PARAMETERS OF THE VIDEO CLASSES OF THE TWO VIDEO SEQUENCES.

Video Classes	Video 2 "Coastguard" *1500 Kbps*				Video 1 "Mobile" *1668 Kbps*			
f_k	f_2	f_3	f_5	f_7	f_2	f_3	f_5	f_7
λ_k (dB/Kbps)	0.0105	0.0064	0.0048	0.0042	0.0170	0.0064	0.0042	0.0031

TABLE II.
Y-PSNR OF THE TWO VIDEO SEQUENCES USING VARIOUS APPROACHES (H_n^I = 40 METERS)

Network Bandwidth		Y-PSNR (dB)		
		AODV/LB	DCS	AFP (1-hop information cell)
Average T =5.5 Mbps	V_1	32.47	35.21	35.61
	V_2	31.70	33.32	33.32

TABLE II.
Y-PSNR OF THE TWO VIDEO SEQUENCES USING VARIOUS APPROACHES (H_n^I = 80 METERS)

Network Bandwidth		Y-PSNR (dB)			
		AODV/LB	DCS	AFP (1-hop information cell)	AFP (2-hop information cell)
Average T =5.5 Mbps	V_1	X	X	28.19	29.80
	V_2	X	X	31.26	31.70
Average T =10 Mbps	V_1	30.47	34.46	35.61	35.61
	V_2	31.92	33.08	33.32	33.32

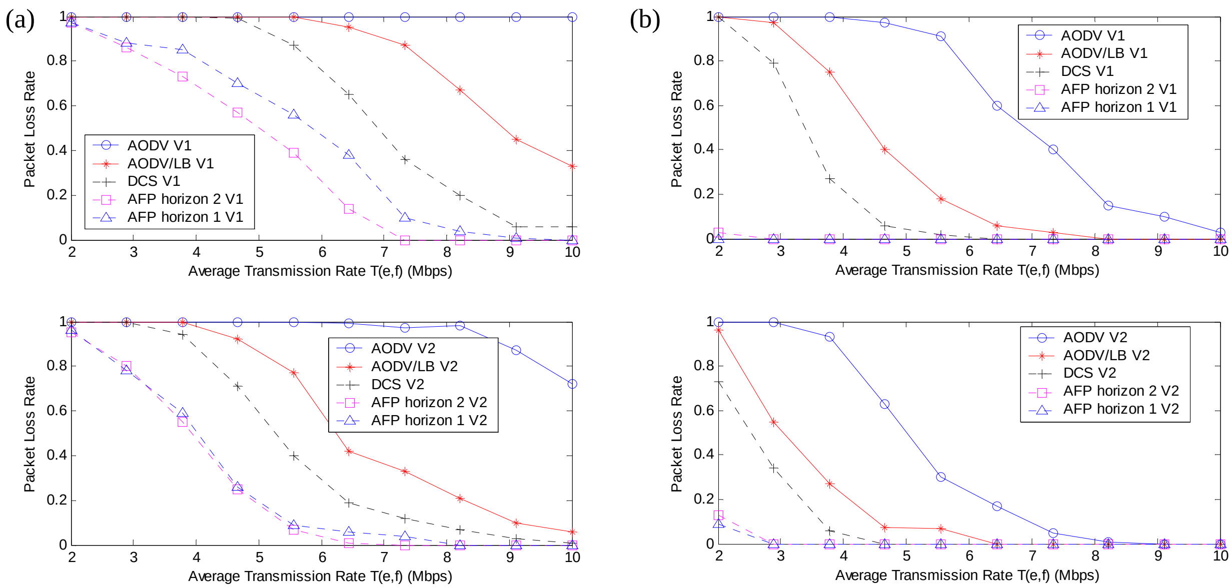

Fig. 10. (a) Packet loss rate vs. average transmission bandwidth using different approaches (H_n^I = 80 meters).
(b) Packet loss rate vs. average transmission bandwidth using different approaches (H_n^I = 40 meters).

C. *Reducing the frequency of learning*

When the interference range is 40 meters, Figure 10(b) shows that the AFP with 1-hop information cell is better than with 2-hop information cell, since 1-hop information cell has smaller cost of information exchange. In addition to reducing the information horizon, reducing the frequency of learning b_n / c at all the nodes can also reduce the cost of information exchange. Figure 11 shows the packet loss rate of the two applications with different information horizon when b_n / c changing from 1 to 0.5. As the learning frequency b_n / c decreases, the packet loss rate decreases with the cost of information exchange. However, it is shown that when $b_n / c <$ 0.6, the AFP becomes inefficient and the packet loss rate starts increasing for both applications. In other words, changing the frequency of learning will also lead to a trade-off between the learning efficiency and the information overhead. The information overhead decreases when the learning frequency b_n / c decreases and

hence, the packet loss rate decreases. However, when the learning frequency is too slow ($b_n / c <$ 0.6), the learning efficiency decreases and this results in an increasing packet loss rate.

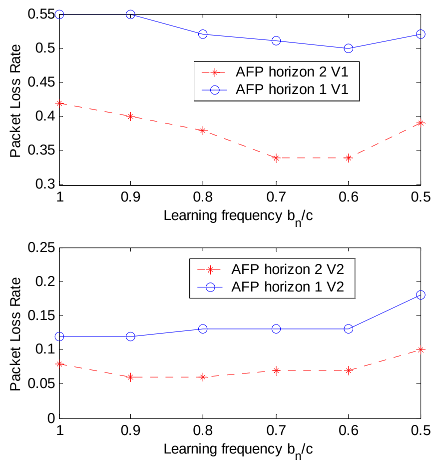

Fig. 11. Packet loss rate vs. learning frequency b_n / c (average T =5.5 Mbps, H_n^I = 80 meters).

D. *Impact of the primary users*

The simulation implies that the reward of information is also impacted by the existence of the primary users. Next, we consider the impact of the primary users, which always have higher priority to access the pre-assigned frequency channels than the network nodes in Figure 8. Assume that the frequency channel F_1 is occupied by the primary users with time fraction ρ =0%, 20%, 40%, 60%, and 80% around a certain congestion region (network nodes n = 7, 11, 12) in Figure 8. Figure 12 shows the packet loss rate for the two video streams using the AFP with various information horizons. The average transmission rate is set to 5.5 Mbps, b_n / c = 1, and the interference rage is 80 meters.

The results show that as the time fraction ρ increases, the packet loss rates of both applications increase, since fewer resources are available for the secondary users to transmit the packets. As the simulation in the previous subsection, when the interference rage is 80 meters, AFP with 2-hop information cell still performs better than 1-hop information cell case. Interestingly, for application V_1, AFP with 3-hop information cell performs even better in a large ρ case, even though more cost of information is needed. This is because the congestion region are more likely to be discovered at the source node n =1 and detour the packets through other routes. However, such advantage is not exploited by the application V_2, since its destination node is affected by the primary users and there is no way to detour the packets. Note that when there is no primary user (ρ = 0),

AFP with 3-hop information cell performs worse than 2-hop case due to the larger cost of information exchange.

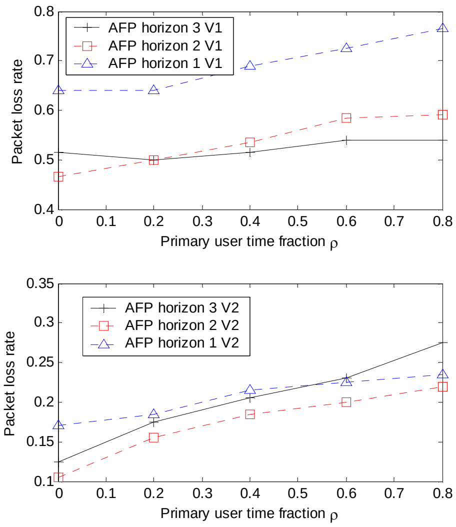

Fig. 12. Packet loss rate vs. time fraction ρ of the primary users occupying frequency channel F_1 around network node n = 7, 11, 12 (average T =5.5Mbps, b_n / c = 1, H_n^I = 80 meters).

E. *Impact of mobility*

In this subsection, we consider the impact of mobility on the video performance. We adopt a well-known mobility model, the "random walk" [26], in which the relay nodes (secondary users) shown in Figure 8 randomly select a direction at each time slot and move at a fixed speed v. We simulate the speed v ranging from 0 to 1 meters/sec. We assume that there is no primary user, i.e. $\rho = 0$. The average transmission rate is set to 8 Mbps, b_n / c = 1, and the interference rage is 80 meters. Figure 13 illustrates the packet loss rate as the mobility changes for different information horizons. The results show that the mobility degrades the performance of both applications. When the mobility v is small, AFP with information horizon $h = 2$ performs better than with information horizon $h = 1$ as in the previous simulations with H_n^I = 80 meters. However, for video V_2, when the mobility exceeds 0.6 meters/sec, the best information horizon changes from $h = 2$ to $h = 1$. This is because the increased mobility will decrease the information accuracy and hence, the required information horizon also decreases. Note that for video V_1, the AFP with information horizon $h = 2$ still performs better than with information horizon $h = 1$. This is because the video V_1 has a longer route and thus, modeling more interfering neighbor nodes, using a larger information horizon, is still beneficial.

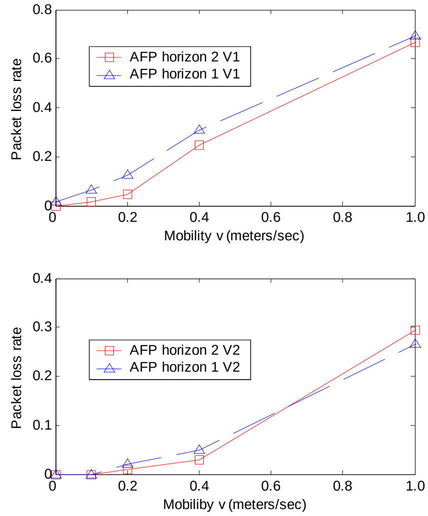

Fig. 13. Packet loss rate vs. mobility v of the secondary users (network relays) (average T = 8Mbps, $\rho = 0$, $b_n / c = 1$, H_n^I = 80 meters).

VIII. CONCLUSIONS

In this paper, we show that the distributed resource management solution using adaptive fictitious play significantly improves the performance of delay sensitive applications transmitted over a multi-hop cognitive radio network. We assume that the autonomous secondary users are able to learn the spectrum opportunities based on the information exchange. The proposed approach can also be used to support QoS for general multi-radio wireless networks, when there is no primary user. This situation is also brought up in [4], when the secondary users are competing in the unlicensed band (i.e. ISM band), where there is no primary user. Importantly, based on the value of the obtained information (i.e. the impact on decreasing the expected end-to-end delay), we define the information horizon in our adaptive fictitious play. In addition to the reward, the cost of the information exchange is also considered in terms of transmission time overheads. Various approaches of decreasing this time overhead are discussed, and their performance impact is quantified.

The information horizon is assumed to be fixed in this paper for different priority classes over the whole wireless networks. However, our simulation results show that the benefit from various information horizons can be different for distinct applications with various delays and quality impacts, especially when primary users are present in the network at different locations. Exploring what are optimal information horizons if the applications and network conditions are changing forms an interesting future research topic in the multi-hop cognitive radio networks.

REFERENCES

[1] Federal Communications Commission (FCC), "Spectrum Policy Task Force," *ET Docket* no. 02-135, Nov 15, 2002.

[2] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," in *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, Feb 2005.

[3] J. Mitola etal., "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13-18, Aug. 1999.

[4] I. F. Akyildiz, W. –Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks: The International Journal of Computer and Telecommunication Networking*, vol. 50, issue 13, Sep 2006.

[5] T. X. Brown, "An analysis of unlicensed device operation in licensed broadcast service bands," in *Proc. IEEE DySPAN 2005*, Nov 2005, pp. 11-29.

[6] C. Cordeiro, K. Challapali, D. Birru and S. Shankar N, "IEEE 802.22: An Introduction to the First Wireless Standard based on Cognitive Radios," *Journal of Communications, Academy Publishers*, vol. 1, issue 1, Apr 2006.

[7] IEEE 802.22 Working Group on Wireless Regional Area Networks, http://www.ieee802.org/22/.

[8] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework" *IEEE Journal on Selected Areas in Communications (JSAC): Special Issue on Adaptive, Spectrum Agile and Cognitive Wireles Networks*, vol. 25, no. 3, pp. 589-600, April, 2007.

[9] D. Niyato and E. Hossain, "A game-theoretic approach to competitive spectrum sharing in cognitive radio networks," in *Proc. IEEE WCNC'07*, Hong Kong, 11-15 March, 2007.

[10] Z. Han, C. Pandana, and K. J. Ray Liu, ``Distributive Opportunistic Spectrum Access for Cognitive Radio using Correlated Equilibrium and No-regret Learning", in *Proceedings of IEEE Wireless Communications and Networking Conference*, 2007.

[11] K.-D. Lee and V.C.M. Leung, "Fair Allocation of Subcarrier and Power in an OFDMA Wireless Mesh Network", *IEEE J. Sel. Areas in Commun.*, vol. 24, no. 11, pp. 2051-2060, Nov. 2006.

[12] H. Wu, F. Yang, K. Tan, J. Chen, Q. Zhang, Z. Zhang, "Distributed Channel Assignment and Routing in Multi-radio Multi-channel Multi-hop Wireless Networks", in *IEEE JSAC special issue on multi-hop wireless mesh networks*, Volume: 24, Issue: 11, page(s): 1972-1983, Nov. 2006.

[13] V. Brik, E. Rozner, S. Banarjee, P. Bahl, "DSAP: a protocol for coordinated spectrum access," in *Proc. IEEE DySPAN 2005, Nov. 2005*, pp. 611-614.

[14] H.-P. Shiang and M. van der Schaar, "Informationally Decentralized Video Streaming over Multi-hop Wireless Networkss," *IEEE Trans. Multimedia*, vol. 9, issue 6, Sep 2007.

[15] D. Fudenberg and D.K. Levine, *The Theory of Learning in Games*, MIT Press, Cambridge, MA, 1998.

[16] H. P. Young, *Strategic Learning and its Limits*, Oxford University Press, NY, 2004.

[17] D. Bertsekas, R. Gallager, *Data Networks*, Prentice Hall, Inc. Upper Saddle River, NJ, 1987.

[18] Y. Shoham, R. Powers, and T. Grenager, "If multi-agent learning is the answer, what is the question?" *Artificial Intelligence*, 2007. to appear.

[19] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in *Proc. ACM Internat. Conf. on Mob. Computing and Networking (MOBICOM)*, 2004, pp. 114-128.

[20] D. Krishnaswamy, "Network-assisted Link Adaptation with Power Control and Channel Reassignment in Wireless Networks," *3G Wireless Conference*, pp. 165-170, 2002.

[21] D. Krishnaswamy, et al. "A cross-layer cross-overlay architecture for proactive adaptive processing in mesh networks," *2nd IEEE Workshop on Wireless Mesh Networks (WiMesh 2006)*, 2006. pp. 74-82.

[22] D. Jurca and P. Frossard, "Media Streaming with Conservative Delay on Variable Rate Channels," in *Proceedings of IEEE international Conference on Multimedia and Expo (ICME, 2006)*, 2006.

[23] C. E. Perkins, E. M. Royer, "Ad hoc on-demand distance vector routing," *in Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications*, pp. 90-100, Feb 1999.

[24] G. D. Kondylis and G. J. Pottie, "Dynamic Channel Allocation Strategies for Wireless Packet Access," *IEEE VTC*, Amsterdam, Sep 1999.

[25] J. Chakareski and P. Frossard, "Rate-Distortion Optimized Distributed Packet Scheduling of Multiple Video Streams

Over Shared Communication Resource," *IEEE Transactions on Multimedia*, vol. 8, no. 2, Apr, 2006.

[26] T. Camp, J. Boleng, V. Davies, "A survey of mobility models for ad hoc network research," in *Wireless Communications and Mobile Computing (WCMC)*, vol. 2, no. 5, pp. 483-502, 2002.